MOBIUS ONE BILLION DOCUMENT BENCHMARK TEST ON AMAZON WEB SERVICES

A DIGITAL WORKPLACE FOR THE MODERN ENTERPRISE

Mobius is a flexible, highly scalable, performance-oriented content services platform that ingests, parses, indexes, archives, and provides access to any type formation, at any time, and anywhere in the enterprise. It is deployable on-premises, in the cloud, or in a hybrid model.

**The purpose of the Mobius One Billion Document Benchmark test was to:**

- Demonstrate the scalability of Mobius in an Amazon Web Services (AWS) Cloud environment

- Demonstrate the reliability of Mobius with ever increasing ingestion volumes

- Demonstrate the performance of the horizontally scalable Mobius architecture

The benchmark test requirements were to:

- Ingest, parse, archive and index 1 Billion documents in the Mobius Repository

- Ingest the documents faster and with less hardware than the competition

- Provide full text indexing of ingested content

- Parse text inputs to identify section and document boundaries

- Exhibit scalable and steady performance through the benchmark test

ARCHITECTURE

The following diagram shows the architecture used to deploy Mobius on AWS EC2 to ingest, parse, archive and full text index 1 Billlion documents, which equaled 100 Terabytes of input data.
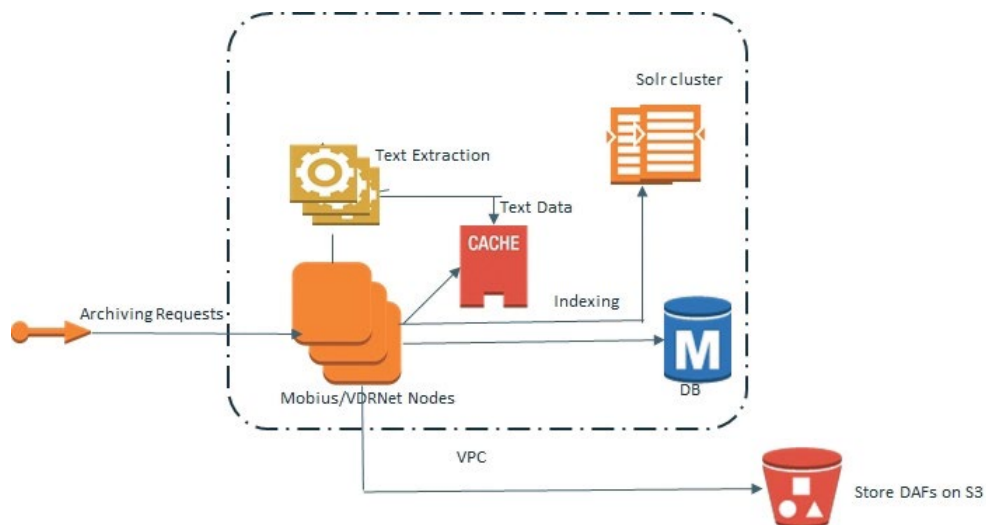
AWS CloudFormation gives developers and systems administrators an easy way to create and manage a collection of related AWS resources, provisioning and updating them in an orderly and predictable fashion. The Amazon EC2 CloudFormation Template permits 1-click deployment of a Mobius Stack:
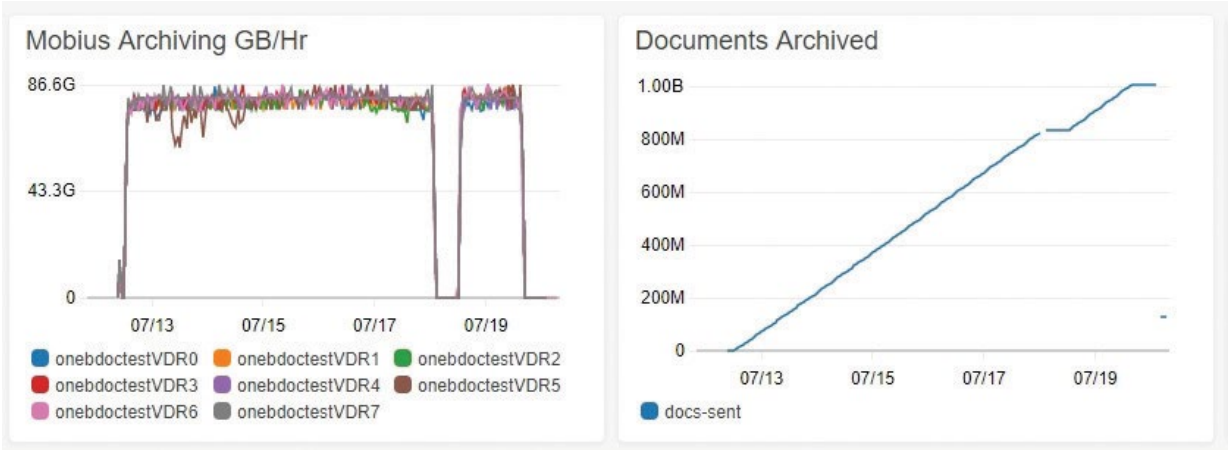
- 8 Mobius ViewDirect Repository EC2 (c3.2xl) instances in an Autoscaling stack

- 18 Apache SolrCloud EC2 (m3.2xl) instances managed by ZooKeeper (m4.l)

- AWS S3 For Archive Storage

- AWS RDS (r3.2xl) for Database

- AWS ElastiCache (r3.4xl)

RESULTS

Mobius ingested, parsed, and archived one billion documents that equaled 100 Terabytes of information. It delivered near real-time full text indexing, with only a 2-minute lag between archiving and full text indexing. The benchmark test was completed in 7 days and 5 hours, and Mobius exhibited steady and scalable performance.
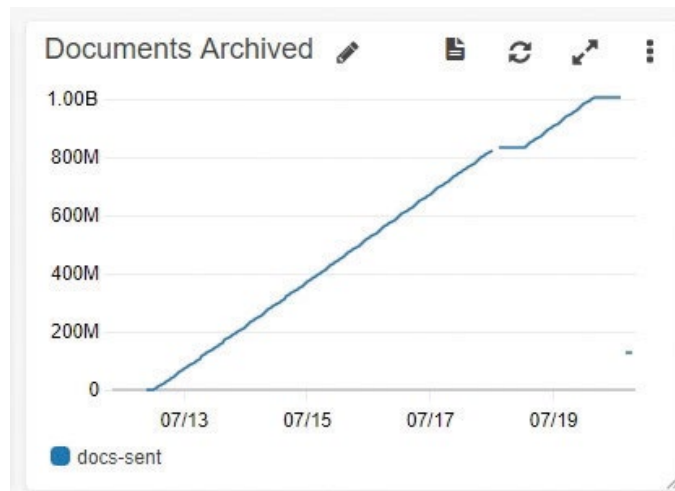
Mobius achieved maximum archiving speed quickly and was stable throughout the 7 days of testing. The AWS EC2 instance suffered a hardware failure during the test. Mobius was not fazed by the failure; archiving was stopped, the failed hardware node was replaced, and archiving continued from where it was stopped.

Mobius Archiving GB/Hr

- onebdoctestVDR0
- onebdoctestVDR1
- onebdoctestVDR2
- onebdoctestVDR3
- onebdoctestVDR4
- onebdoctestVDR5
- onebdoctestVDR6
- onebdoctestVDR7

Documents Archived

- docs-sent

Mobius compression of documents for archival resulted in a 40% reduction in disk utilization, which provides additional storage cost savings on AWS S3. Performance was not hindered using standard Mobius archive compression.

Mobius DAF Storage

- BucketSizeBytes
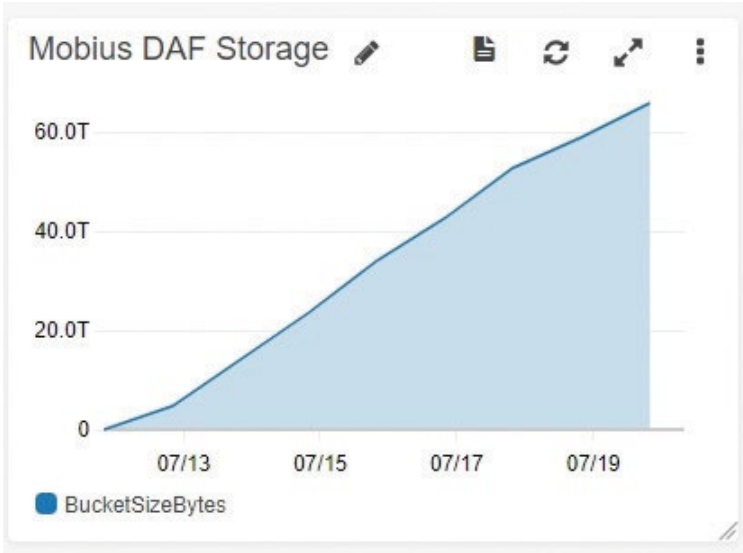
The Mobius node CPU Utilization was less than 50% when using standard Mobius archive compression. Additional reduction in disk utilization without performance degradation is possible with enhanced Mobius archive compression.
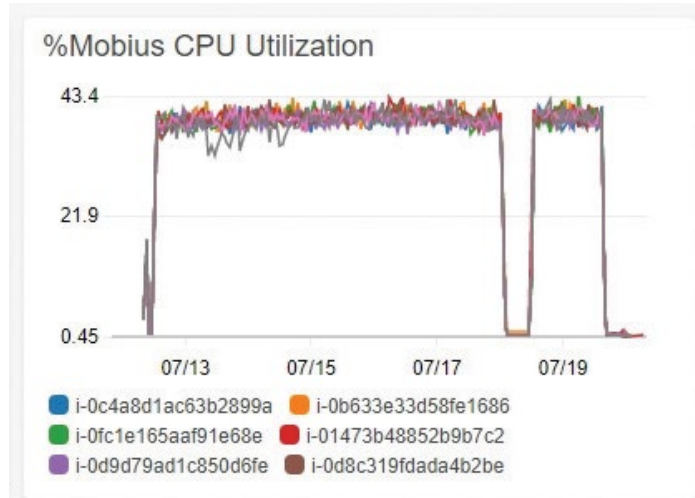
Mobius Database and Full Text Indexing CPU Utilization peaked at 10%, leaving 90% of CPU capacity unused. Additional cost savings can be achieved using a lower class of machines and increasing CPU utilization without impacting performance. Additionally, it was observed that memory utilization was lower than expected.
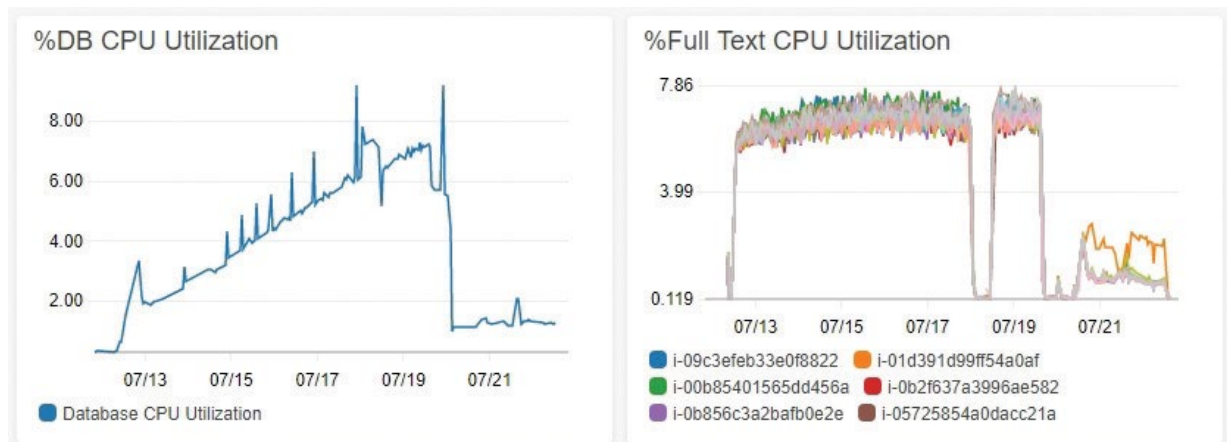
SUMMARY

The scalability and configurability of Mobius lets customers right size deployments to optimize multiple enterprise goals. In comparison to the competition's billion document deployment on AWS, the Mobius One Billion Document Benchmark test achieved a 250% faster completion time while running on 10% fewer instances of the same hardware, with low CPU utilization on these instances. This resulted in a 30% reduction in cost for the benchmark, with opportunities to further reduce storage and deployment costs. This shows how Mobius deployments can be easily configured to match specific enterprise operational goals for runtime cost, fault tolerance and operational security without impeding performance and scalability.

ASG Technologies is a global software company providing the only integrated platform and flexible end to end solution for the information powered enterprise. ASG is the only solutions provider for both Information Management and IT Systems and has over 3,000 customers worldwide. To learn more visit www.asg.com.

ASG_Billion_Benchmark_20191015